

SCREENING METHOD FOR GENE VARIATION

[0001]

BACKGROUND OF INVENTION

Field of the Invention

5 The present invention relates to a variation detection method and a system for it useful for screening of a gene variation etc.

[0002]

Related Background Art

10 One of the techniques for sequencing nucleic acid etc. or for detecting the sequence is to utilize a DNA array. USP 5,445,934 discloses a DNA array where 100,000 or more oligonucleotide probes are bonded in 1 inch square. Such a DNA array has an advantage that
15 many characteristics can be examined at the same time with a very small sample amount. When a fluorescence-labeled sample is poured onto such a DNA chip, DNA fragments in the sample bind to probes having a complementary sequence fixed on the DNA chip, and only
20 that part can be discriminated by fluorescence to elucidate the sequence of the DNA fragment in the DNA sample.

[0003]

25 Sequencing By Hybridization (SBH) is a method for examining the base sequence utilizing such a DNA array and the details are described in USP 5,202,231. In the SBH method, all possible sequences of an

TOP SECRET

oligonucleotide of a certain length are arranged on the substrate, then fully matched hybrids formed by hybridization reaction between probes and the sample DNA are detected. If a set of fully matched hybrids is obtained, the set will give an assembly of overlapping sequences with one base shift being a part of one certain sequence, which sequence is extracted for calling.

[0004]

Not only the SBH method, when complementariness between an oligonucleotide and a sample DNA is examined, it is very difficult to call whether a hybrid was formed or not using one probe for one test item, since the stability of a hybrid differs sequence to sequence, and there is no perfect signal for calling the full complementariness. Science vol. 274 p.610-614, 1996 discloses a method for calling comparing the signal intensity of a perfect match hybrid and the weaker intensities of one-base mismatch hybrids. In this method, 15-mer oligonucleotide probes differing from each other only by one mismatching base at the center of the sequence are prepared, and the fluorescence intensities of the hybrids of the probes are compared, and when the intensity of the full matched hybrid is stronger than that of other hybrids by a predetermined rate, it is called positive.

[0005]

FOIA b 7 - D

Further, USP 5,733,729 discloses a method using a computer to know a base sequence of a sample from a comparison of fluorescence intensities of obtained hybrids for more accurate calling.

5 [0006]

However, actual binding strength of a hybrid depends on the GC content etc., and difference of the fluorescence intensity between a full match hybrid and a one-base mismatch hybrid also varies in a considerable range depending on the sequence. Thus, a method for calling whether a sequence is fully complementary to a probe or not, using a 15 mer oligonucleotide probe to compare it with other three probes having one mismatched base at the center thereof, can provide more accuracy if each stability is evaluated theoretically or empirically before comparison.

[0007]

In addition, accurate calling requires precise quantification of signals, and therefore, precision apparatuses such as a confocal laser microscope. Furthermore, in order to measure the fluorescence intensity of a hybrid of every probe and to determine the gene sequence by analyzing the data, a large-scale computer apparatus as well as a detection apparatus for reading the arrays are further required. Therefore this is a big obstacle for ready use of the DNA array.

[0008]

On the other hand, gene diagnosis using such a DNA array may be used in group medical examination, individual gene examination or gene-polymorphism study.

5 In such a case, however, the above described precise measurement and analysis are not always required, where a large amount of samples are rapidly treated at a low cost in order to find out variated samples concerning a specific item from a large number of normal samples.

10 Further, the precision apparatus and analysis as described above will cost much. Accordingly, a concept that screening of the presence or absence of a variation is first performed, and then, detailed examinations are carried out about the samples
15 suspected of variation by the screening will save time and cost.

[0009]

SUMMARY OF THE INVENTION

One object of the present invention is to provide
20 a method suitable for mass screening so as to determine rapidly the presence or absence of a gene variation without need of an expensive apparatus and a complex analysis.

[0010]

25 The present invention provides a DNA array in which a group of probes which will give strong signals forming hybrids with a normal gene sequence, and a

FOR PUBLICATION

group of probes having sequences expected to form hybrids with gene variants are separately arranged, on the premise that the base sequence of a normal gene and those of variants have already been established.

5 Furthermore, the above described object is achieved by providing a detection method using such an array.

[0011]

According to one aspect of the present invention, there is provided a method for screening of the presence or absence of variation in a region of a nucleic acid comprising the steps of:

(a) preparing a test nucleic acid corresponding to the region;

(b) preparing a probe having a base sequence fully complementary to a normal sequence of the region, and a plurality of probes each having at least one base not complementary to the normal sequence;

(c) fixing the probes in separate regions on a surface of a substrate to prepare a DNA array substrate;

(d) reacting the test nucleic acid with the probes on the DNA array substrate;

(e) measuring signals in each region totaly where the signals are originated from respective hybrids formed between the test nucleic acid and one of the probes; and

(f) determining the presence or absence of

mutation in the test nucleic acid comparing with a histogram pattern of signals of all regions obtained using a normal sample without variation.

[0012]

5 According to another aspect of the invention, there is provided a DNA array substrate for screening a variation in a region of a nucleic acid, wherein

 a full match probe fully complementary to a normal sequence of the region, and a plurality of mismatch
10 probes having at least one base mismatch to the sequence are arranged on the substrate; and

 the probes are arranged to form at least two separate regions selected from:

 a first region containing at least one probe which
15 provides a signal of a certain intensity on reaction with a nucleic acid having the normal sequence,

 a second region containing at least one probe which provides a weaker signal than the probe of the first region on reaction with a nucleic acid having
20 normal sequence, and

 the third region containing at least one probe which provides no signal on reaction with a nucleic acid having normal sequence.

[0013]

25 According to still another aspect of the present invention, there is provided a system for detecting variation comprising a DNA array substrate as described

above and a signal measuring apparatus which measures signals from separate regions of the DNA array substrate.

[0014]

5 The present invention can provide a method suitable for mass screening, so as to rapidly determine only the presence or absence of a gene variant, without need of an expensive apparatus and complex analysis.

[0015]

10 BRIEF DESCRIPTION OF THE DRAWINGS

 FIG. 1A is a schematic example of arrangement of separate regions on a DNA array substrate, and FIG. 1B shows a histogram of fluorescence intensities (pattern) corresponding to separate regions of the above arrangement, measured using a line sensor:

15

 FIG. 2 shows an arrangement of 64 types of DNA probes;

 FIG. 3 shows a pattern of the fluorescence intensities obtained by hybridization;

20

 FIG. 4 shows normal fluorescence intensities obtained by hybridization where Gr. NO. denotes group number;

25

 FIG. 5 is a distribution pattern of fluorescence intensities corresponding to the separate regions of the arrangement, measured using a line sensor;

 FIG. 6 is a distribution pattern of fluorescence intensities corresponding to the separate regions of

the arrangement, measured using a line sensor;

FIG. 7 is a distribution pattern of fluorescence intensities corresponding to the separate regions of the arrangement, measured using a line sensor;

5 FIG. 8 is a distribution pattern of fluorescence intensities corresponding to the separate regions of the arrangement, measured using a line sensor; and

FIG. 9 shows normal fluorescence intensities obtained by hybridization where Gr. NO. denotes group number.

[0016]

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a screening method for gene variants using a DNA array in which a group of probes which will give strong signals forming hybrids with a normal gene sequence, and a group of probes having sequences expected to form hybrids with gene variants are separately arranged, on the premise that the base sequence of a normal gene and those of variants have already been established. Furthermore, the above described object is achieved by providing a detection method using such an array.

[0017]

Here, the present invention will be described in detail with examples where signals from the separate regions are fluorescence. However, signals in the present invention are not limited to fluorescence but

TOP SECRET

may be other light signals or electric signals.

Binding strength between single-stranded nucleic acids to form a hybrid is controlled by various factors, and when a probe having a length of about 12
5 mer to 25 mer is used, it is practically difficult to perfectly exclude hybrids having one-base mismatches.

[0018]

When the signal of the hybrid to be detected is a light such as fluorescence, the following phenomena are
10 observed. Fluorescence stability of a hybrid having two mismatches (two-base mismatch hybrids) is much lower than that of the one-base mismatched hybrid, regardless of positions, continuity or discontinuity of the mismatched bases. On the other hand, signal from a
15 three-base mismatch hybrid is hardly observed.

However, one-base mismatch hybrids may have more than 50% signal intensity of a full match hybrid. Thus, when a sample is a nucleic acid of normal sequence, strong fluorescence is observed in a region where a
20 full match probe and probes having one mismatching base have been arranged, while the fluorescence intensity in a region where hybrids of low stability are formed is almost zero. On the other hand, when a sample is a nucleic acid of a variant sequence, the probe fully
25 complementary to the normal sample makes a mismatch hybrids with the sample. Thus the fluorescence is weaker at a region than that in case of the normal

FOR SEARCH

sample, at the same time, fluorescence from a full match hybrid and one-base mismatch hybrids with the sample appears in another region where low signals are expected in case of the normal nucleic acid.

5 Accordingly, by comparing fluorescence histogram of the regions, it is possible to distinguish between the normal nucleic acid and the variant nucleic acid.

[0019]

10 In the present invention, a DNA array substrate where probes are arranged in separate regions according to the fluorescence intensities of their hybrids with a normal nucleic acid is used for more accurate calling. First, hybridization reaction of the normal nucleic acid with each probe is performed, and based on the
15 fluorescence intensities of the hybrids obtained, separate regions each containing probes corresponding to strong fluorescence, no fluorescence, and moderate (weak) fluorescence are located at predetermined positions on the substrate.

20 [0020]

When performing a test on multiple items simultaneously, the substrate should be divided into areas for respective items. In each area, probes having full match and one-base mismatch sequences to
25 the normal gene sequence are arranged in a region, which are expected to give high signals, and the other probes having more than two-base mismatch are arranged

FOIA b 7 - D

in a separate region depending on the items. Thus we can discriminate normal test samples from varied ones.

[0021]

5 The arrangement can be determined according to the type of the sensor to be used. For example, when a line sensor is used, each region is arranged from the left to right in the substrate in order of the strength of fluorescence obtainable by hybridization with the normal nucleic acid. Thus, the fluorescence intensity will be maximum in the left, then gradually decrease and come to zero in several regions in the right of the substrate. When an area sensor is used, it is necessary to evaluate the fluorescence quantity of at least two separate regions containing a group of probes which will provide the maximum fluorescence and a group of probes which will not provide any fluorescence respectively.

[0022]

20 This will be described more specifically.

[0023]

In the present invention, for example, whether a gene is normal or not can be called by providing a region containing a probe that forms a hybrid with a normal nucleic acid and the other region containing probes that form hybrids with variant genes separately on a DNA array, and taking a ratio of the signal from a

hybrid corresponding to the normal nucleic acid and to the signals from hybrids corresponding to variant genes.

[0024]

5 However, because one-base mismatched hybrids sometimes have strong signals, it is difficult to detect variation in the test sample when a DNA array where only probes being full match or one-base mismatch to a normal sequence are arranged is used. It is
10 important that probes having two-base mismatch to the normal nucleic acid is present in the DNA array, which might be one-base mismatch probes to the test sample to form hybrids having strong signals.

[0025]

15 Therefore, in the present invention, in order to judge more accurately, a preferable method is as follows. Probes having full match and one-base mismatch sequences to the normal nucleic acid sequence, which most samples in mass screening have, are arranged
20 in a specified region on a DNA array substrate. In addition, probes having two- or three-base mismatch are arranged in a region different from the above region for high stability hybrids. Then, a hybridization reaction is performed with the normal nucleic acid or a
25 sample, using the DNA array substrate of such an arrangement. Then the total of signals for each separate region is measured, and a pattern obtained

with a sample nucleic acid is compared that with the normal nucleic acid are compared to determine the presence or absence of variation.

[0026]

5 For example, when 64 probes ($4 \times 4 \times 4 = 64$) where different bases are arranged at three positions in 18 base length (Table 1, 5'-terminus is on the left) are used supposing that variation occurs at these three points, for any sample there should be present a probe
10 fully complementary to the sample, nine one-base mismatch probes, 27 two-base mismatch probes and 27 three-base mismatch probes. Furthermore, it is relatively easy to set conditions of the hybridization reaction such that fluorescence is observed with the
15 full match and one-base mismatch hybrids but not with three-base mismatch hybrids. Some two-base mismatch probes form hybrids and others not, depending to their sequences.

[0027]

[Table 1]

SEQ ID NO:	Sequence	SEQ ID NO:	Sequence
1	GATGGGACTCAAGTTCAT	33	GATGGGCCTCAAGTTCAT
2	GATGGGACTCAGGTTTCAT	34	GATGGGCCTCAGGTTTCAT
3	GATGGGACTCACGTTTCAT	35	GATGGGCCTCACGTTTCAT
4	GATGGGACTCATGTTTCAT	36	GATGGGCCTCATGTTTCAT
5	GATGGGACTCGAGTTTCAT	37	GATGGGCCTCGAGTTTCAT
6	GATGGGACTCGGGTTTCAT	38	GATGGGCCTCGGGTTTCAT
7	GATGGGACTCGCGTTTCAT	39	GATGGGCCTCGCGTTTCAT
8	GATGGGACTCGTGTTTCAT	40	GATGGGCCTCGTGTTTCAT
9	GATGGGACTCCAGTTTCAT	41	GATGGGCCTCCAGTTTCAT
10	GATGGGACTCCGGTTTCAT	42	GATGGGCCTCCGGTTTCAT
11	GATGGGACTCCCGTTTCAT	43	GATGGGCCTCCCGTTTCAT
12	GATGGGACTCCTGTTTCAT	44	GATGGGCCTCCTGTTTCAT
13	GATGGGACTCTAGTTTCAT	45	GATGGGCCTCTAGTTTCAT
14	GATGGGACTCTGGTTTCAT	46	GATGGGCCTCTGGTTTCAT
15	GATGGGACTCTCGTTTCAT	47	GATGGGCCTCTCGTTTCAT
16	GATGGGACTCTTGTTTCAT	48	GATGGGCCTCTTGTTTCAT
17	GATGGGGCTCAAGTTCAT	49	GATGGGTCTCAAGTTCAT
18	GATGGGGCTCAGGTTTCAT	50	GATGGGTCTCAGGTTTCAT
19	GATGGGGCTCACGTTTCAT	51	GATGGGTCTCACGTTTCAT
20	GATGGGGCTCATGTTTCAT	52	GATGGGTCTCATGTTTCAT
21	GATGGGGCTCGAGTTTCAT	53	GATGGGTCTCGAGTTTCAT
22	GATGGGGCTCGGGTTTCAT	54	GATGGGTCTCGGGTTTCAT
23	GATGGGGCTCGCGTTTCAT	55	GATGGGTCTCGCGTTTCAT
24	GATGGGGCTCGTGTTTCAT	56	GATGGGTCTCGTGTTTCAT
25	GATGGGGCTCCAGTTTCAT	57	GATGGGTCTCCAGTTTCAT
26	GATGGGGCTCCGGTTTCAT	58	GATGGGTCTCCGGTTTCAT
27	GATGGGGCTCCCGTTTCAT	59	GATGGGTCTCCCGTTTCAT
28	GATGGGGCTCCTGTTTCAT	60	GATGGGTCTCCTGTTTCAT
29	GATGGGGCTCTAGTTTCAT	61	GATGGGTCTCTAGTTTCAT
30	GATGGGGCTCTGGTTTCAT	62	GATGGGTCTCTGGTTTCAT
31	GATGGGGCTCTCGTTTCAT	63	GATGGGTCTCTCGTTTCAT
32	GATGGGGCTCTTGTTTCAT	64	GATGGGTCTCTTGTTTCAT

bioRxiv preprint doi: <https://doi.org/10.1101/014650>; this version posted June 12, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[0028]

When these 64 probes are grouped into every eight probes in order of the fluorescence intensity obtained by hybridization with the normal nucleic acid (hereinafter "fluorescence intensity of a probe(s)" means expected intensity of a hybrid of the probe with a nucleic acid of normal sequence, if not otherwise stated), the total fluorescence intensity of the first group should be extremely high and the total fluorescence quantity of the sixth, seventh and eighth groups should be zero. Such classification by the fluorescence intensity may be performed empirically or theoretically through calculation.

[0029]

Then, the hybridization reaction is performed in an actual system and the total quantity of fluorescence is determined for each classified group. Particularly, the fluorescence quantities of the sixth, seventh and eighth groups are important. Normally, fluorescence is not expected for these groups. When fluorescence is detected unexpectedly for these groups, it is understood that the sample is not normal but having variation at a part of the sequence. When the fluorescence is not detected for these groups, most of the sequences are normal.

[0030]

For more accuracy, however, it is necessary to

compare the fluorescence quantity of the first group with a normal case. When the fluorescence quantity is significantly lower than that expected for normal sequence, base variation is suspected. Further the histogram of the fluorescence intensity in the regions having medium fluorescence intensity, i.e. the second, third and fourth groups, should be compared with that of the normal one.

[0031]

For more effective detection of variation, one may consider a method wherein the probes are arranged in order of signal intensity, that is, lined in the order of sequences of full match, one-base mismatch, two-base mismatch, and three-base mismatch to the normal sequence from one end, so that signal intensity may be zero in the region opposite to the region where the full match probe is placed. In this case, the total distribution pattern of the fluorescence intensity is examined. For example, when using an arrangement of the separate regions shown in FIG. 1A (one column corresponds to one separate region), a monotonously decreasing pattern as shown in FIG. 1B is obtained for the normal case while patterns having peaks in the intermediate regions will be obtained with variants.

[0032]

The above arrangement method is similarly used for the case where the probes for multi-item testing are

TOP SECRET

arranged on the same substrate. For each item, first,
the full matched probe (to the normal sequence), then
one-base mismatched sequences, two-base mismatched
sequences and so on are placed in order of the strength
5 of fluorescence intensity expected.

[0033]

Such concept is universally applicable to any
number of variation, not limited to the above method
where the variation for only three bases is tested.

10 [0034]

In addition, while we explained the cases where
the signals can be obtained when the hybrids are
formed, the method may be set such that signals are not
obtained when the hybrids are formed, and obtained when
15 the hybrids are not formed, depending on the signal
generating system.

[0035]

The sample nucleic acid can be prepared by
extracting from the gene to be tested according to the
20 necessity. The control normal nucleic acid can be
synthesized on the basis of the known sequence.

[0036]

The length of the probe fixed to the substrate is
not limited so long as it is suitable for detection,
25 for example, preferably 8 mer to 30 mer, more
preferably 12 mer to 25 mer.

[0037]

Probes may be fixed to the substrate by various methods, but droplet application by the ink jet method is preferably used in order to arrange spots of fixed probes with high efficiency, high density and high speed.

5

Each spot in the separate region is preferably 70 to 100 μm in diameter and spaced not to connect each other.

[0038]

10

The spot number in each region is set so that the spots expected to have high fluorescence intensity can be measured together and the spots expected to have no fluorescence can be measured together, considering the constitution of the sensor.

15

[0039]

20

A system for detecting variation of the present invention comprises a DNA array substrate wherein plural separate regions are arranged in a prescribed arrangement, and a sensor for measuring the signals in the separate regions subjected to measurement. In addition, for calling the presence or absence of variation using the DNA array substrate where the plural separate regions are provided, the signals from all separate regions may be detected and the results are used for calling the presence or absence of variation or, as a simpler method, signals may be compared between certain regions selected so as to

25

441130 333440

detect the presence or absence of variation.

[0040]

Computer analysis is conveniently performed connecting a computer system to the detection system.

5 For many variants of a gene sequence, histogram patterns of signal intensity are prepared and stored in the computer, which helps the determination of variant easily and correctly.

[0041]

10 As the sensor for signal detection, different types of photodiodes (e.g. Hamamatsu Photonics K.K.-made) are used. For example, some of the divided type silicon photodiode arrays can be used as both line sensor and area sensor. Particularly, those having a
15 photo-receiving face of 1 to 2 mm × (2 to 3) mm are suitably used.

[0042]

Examples

The present invention will be described in more
20 detail referring to Examples below. Herein, "%" means "weight %".

Example 1

Preparation of DNA array for detection of variant gene for line sensor

25 1) Preparation of DNA array linked with 64 types of probes

(1) Probe Design

It is well known that in the base sequence CGGAGG corresponding to the AA248 and AA249 of the tumor suppressor gene p53, frequently observed variation is those the first C to T, the second A to G for AA248, and the third G to T for AA249. Accordingly, aiming at these three positions, 64 types of probes were designed.

[0043]

That is, the designed nucleic acid are 18-mer nucleic acids harboring variegated above mentioned six bases sandwiched between the common sequences, represented by 5'ATGAACNNGAGNCCCATC3' where N corresponds to any of 4 bases, A, G, C and T. Actual probes to detect the above sequence should be have a complementary sequence of 5'GATGGGCTCCGGTTCAT3'.

[0044]

FIG. 2 shows an arrangement of 64 types of DNA probes on a substrate. A sequence 5'ATGAACCGGAGGCCCATC3' corresponding to the normal gene is expected to form a hybrid with DNA of probe 42 of 5'GATGGGCTCCGGTTCAT3' positioned at the third from the right and the third from the top.

(2) Preparation of substrate introduced with maleimide group

Substrate Cleaning

A glass plate of 1 inch square was placed in a rack and soaked in an ultrasonic cleaning detergent

overnight. Then, after 20 min of ultrasonic cleaning, the detergent was removed by washing with water. After rinsing the plate with distilled water, ultrasonic treatment was repeated in a container filled with
5 distilled water, for additional 20 min. Then the plate was soaked in a prewarmed 1N sodium hydroxide solution for 10 min, washed with water and then distilled water.
Surface treatment

Then the plate was soaked in an aqueous solution
10 of 1% silane coupling agent (product of Shin-Etsu Chemical Industry: Trade name KBM 603) at a room temperature for 20 min, thereafter nitrogen gas was blown on the both sides blowing off water to dryness. The silane coupling treatment was completed by baking
15 the plate in an oven at 120°C for 1 hour.
Subsequently, 2.7 mg of EMCS (N-(6-maleimidocaproyloxy) succinimide: Dojin Company) was weighed and dissolved in a 1 : 1 solution of DMSO/ethanol (final
concentration: 0.3 mg/ml). The glass substrate treated
20 with the silane coupling agent was soaked in this EMCS solution for 2 hours to react the amino group of the silane coupling agent with the succimide group of EMCS. At this stage, the maleimide group of EMCS is transferred to the glass surface. After that, the
25 glass plate was washed with ethanol, and dried with nitrogen gas to be used for a coupling reaction with DNA.

with normal p53 sequence

(1) Hybridization reaction

- Blocking reaction

After completion of the reaction, the substrate
5 was washed with a 1 M NaCl/50 mM phosphate buffer
solution (pH 7.0) to wash out thoroughly the DNA
solution on the glass surface. Then, this was soaked
in an aqueous solution of 2% bovine serum albumin and
allowed to stand for 2 hours to carry out a blocking
10 reaction.

- Preparation of model sample DNA

Rhodamine labeled DNA No. 1 (SEQ ID NO: 65) of the
same length as the probes but having the normal
sequence of p53 gene was prepared. The sequence is
15 shown below and rhodamine is bonded to the 5'
terminus. (Synthesis of model sample of DNA)

The labeled DNA No. 65 (single strand) having the
normal sequence of p53 gene (complementary to No. 42)
and the same length in the same region as the probe DNA
20 was prepared. The sequence is as shown below where
rhodamine (Rho) is bound to the 5'-terminal.

No. 65: 5'-Rho-ATGAACCGGAGGCCCATC-3'

-Reaction condition of hybridization

Two milliliters of 50 nM DNA solution of a model
25 sample containing 100 mM NaCl was placed into a
container containing the DNA array substrate for a
hybridization reaction. Initially it was heated at

70°C for 30 min, then the temperature of the incubator was lowered to 40°C and the reaction was continue for 3 hours.

(2) Detection

5 - Method

The detection was performed by connecting an image analysis processing apparatus, ARGUS (a product of Hamamatsu Photonics) to a fluorescence microscope (a product of Nikon).

10 (3) Result

Distribution of the fluorescence quantity on the substrate obtained is shown in FIG. 3.

[0047]

Example 2

15 Preparation of DNA array for detection of variant gene for line sensor

(1) Preparation of DNA array for detection of variation

In Tables 2 and 3, these 64 probes are grouped in every 8 probes in order of intensity based on the above described results. The fluorescence intensity of the first group should be extremely strong, and the total fluorescence quantity of the sixth, seventh and eighth groups should be zero.

[0048]

FOR SECRETION

[Table 2]

Group No.	SEQ ID NO:	Type of Mismatch
1	42	Full mismatch
	58	One-base mismatch
	34	One-base mismatch
	41	One-base mismatch
	46	One-base mismatch
	10	One-base mismatch
	44	One-base mismatch
	43	One-base mismatch
2	26	One-base mismatch
	38	One-base mismatch
	50	Two-base mismatch
	2	Two-base mismatch
	6	Two-base mismatch
	9	Two-base mismatch
	11	Two-base mismatch
	12	Two-base mismatch
3	14	Two-base mismatch
	18	Two-base mismatch
	22	Two-base mismatch
	25	Two-base mismatch
	27	Two-base mismatch
	28	Two-base mismatch
	30	Two-base mismatch
	33	Two-base mismatch
4	35	Two-base mismatch
	36	Two-base mismatch
	37	Two-base mismatch
	39	Two-base mismatch
	40	Two-base mismatch
	45	Two-base mismatch
	47	Two-base mismatch
	48	Two-base mismatch

44150-336400

[0049]

[Table 3]

Group No.	SEQ ID NO:	Type of Mismatch
5	54	Two-base mismatch
	57	Two-base mismatch
	59	Two-base mismatch
	60	Two-base mismatch
	62	Two-base mismatch
	1	Three-base mismatch
	3	Three-base mismatch
	4	Three-base mismatch
6	5	Three-base mismatch
	7	Three-base mismatch
	8	Three-base mismatch
	13	Three-base mismatch
	15	Three-base mismatch
	16	Three-base mismatch
	17	Three-base mismatch
	19	Three-base mismatch

10150 834650

[Table 3 (continued)]

Group No.	SEQ ID NO:	Type of Mismatch
7	20	Three-base mismatch
	21	Three-base mismatch
	23	Three-base mismatch
	24	Three-base mismatch
	29	Three-base mismatch
	31	Three-base mismatch
	32	Three-base mismatch
	49	Three-base mismatch
8	51	Three-base mismatch
	52	Three-base mismatch
	53	Three-base mismatch
	55	Three-base mismatch
	56	Three-base mismatch
	61	Three-base mismatch
	63	Three-base mismatch
	64	Three-base mismatch

[0050]

Then, the surface of the substrate was separated into eight columns to arrange the first group, the second group and so on in order of intensity from the left to the right. Then, the probes were arranged at

10750-884459

the positions of the corresponding probe numbers as shown in FIG. 4 using the ink jet method in the same manner is in Example 1 to prepare the DNA arrays for detecting variation. The lines composed of each group have intervals of 200 μm .

(2) Testing normal gene using DNA array

The hybridization reaction was carried out under the same conditions as in Example 1. Thereafter, the total fluorescence of each group was detected using a line sensor (S 272102: Hamamatsu Photonics K.K.-made). [0051]

The results are shown in FIGS. 4 and 5. The fluorescence of the first group was the highest and the intensity of the second group decreases below the half of that. Fluorescence was hardly observed in the sixth, seventh and eighth groups.

Example 3

Detection of variant gene using DNA array (1)

1. Synthesis of model variant DNA

The labeled DNA No. 66 having the same length as the probes and a sequence complementary to No. 46 probe that differs by one base from the normal sequence of p53 gene was prepared. The sequence is shown below. Rhodamine was bound to the 5'-terminal. The underlined part is the variant position.

No. 66: 5'-Rho-ATGAACCAGAGGCCCATC-3' (SEQ ID NO:66)

Reaction conditions for hybridization

The hybridization reaction was carried out under the same conditions as in Example 1. The concentration of the model sample DNA was 50 nM.

5 Detection by line sensor

After the hybridization reaction, the DNA array was evaluated in the same manner as in Example 2 using a line sensor. As shown in FIG. 6, the fluorescence was observed in the sixth, seventh and eighth groups
10 where the fluorescence was not observed with the normal gene. Thus, it was easily judged that the sample gene was not normal. Further, since the fluorescence intensity was high at the first and the fourth groups corresponding to probe Nos. 35 to 48, it is inferred
15 that they have the variation in the upper right quarter of FIG. 2 with a sequence very close to the normal gene.

[0052]

This result shows that the screening method for
20 variant genes of the present invention is extremely effective.

[0053]

Example 4

Detection of variant gene using DNA array (2)

25 The hybridization reaction was carried out under the same conditions as in Example 3, except that the concentration of the model target gene used for the

hybridization reaction was changed to 5 nM. The result is shown in FIG. 7.

[0054]

5 Since the results similar to Example 3 were obtained to show that the method of the present invention works not depending on the hybridization reaction conditions.

[0055]

Example 5

10 Detection of variant gene using DNA array (3)

Synthesis of variant model sample DNA

15 The labeled DNA No. 67 having the same length as the probes and a sequence complementary to No. 10 probe that differs by one base from the normal sequence of p53 gene was prepared. The sequence is as shown below where Rho represents rhodamine bound to the 5'-terminus. The underlined part is the variant position.

No. 67: 5'-Rho-ATGAACCGGAGTCCCATC-3' (SEQ ID NO: 67)

20 Hybridization reaction

The hybridization reaction was carried out under the same conditions as in Example 1 using the above variation model sample DNA. The concentration of the sample was 50 nM. The result is shown in FIG. 8.

25 [0056]

Fluorescence was observed in the sixth and seventh groups which was not observed for the normal gene,

showing that this was not a normal gene. Since the
fluorescence in the second group was higher than in the
first group which would be the highest for the normal
gene, it is presumed that it has a variation included
5 in the probe sequences of the second group (the upper
left quarter in FIG. 2).

[0057]

This result shows that the screening method for
variant genes of the present invention is feasible not
10 depending on the variation types.

[0058]

Example 6

Preparation of DNA array for area sensor

The probes grouped as shown in Example 2 were
15 arranged by the ink jet method in the separate regions
on the substrate as shown in FIG. 9 to prepare the DNA
array for area sensor.

[0059]

Then, the hybridization reaction was performed
20 using the variant model sample used in Example 4. As a
result, the fluorescence was observed in the sixth,
seventh and eighth groups which was not observed for
the normal gene, therefore, it was easily judged that
this gene was a variant one.

SEQUENCE LISTING

<110>Canon INC.

<120>Screening method for gene variation

<130>CFO 15717

<150>JP 2000-263396

<160>67

<210>1

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>1

gatgggactc aagttcat

<210>2

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>2

gatgggactc aggttcat

<210>3

<211>18

<212>DNA

<213>Artificial sequence

<220>

10.0.0.0

<210>7

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>7

gatgggactc gcgttcat

<210>8

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>8

gatgggactc gtgttcat

<210>9

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>9

gatgggactc cagttcat

<210>10

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>10

gatgggactc cggttcat

<210>11

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>11

gatgggactc ccgttcat

<210>12

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>12

gatgggactc ctgttcat

<210>13

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

U.F. 30-854450

<400>13

gatgggactc tagttcat

<210>14

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>14

gatgggactc tggttcat

<210>15

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>15

gatgggactc tcgttcat

<210>16

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>16

gatgggactc ttgttcat

<210>17

FOR SEQUENCE

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>17

gatggggctc aagttcat

<210>18

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>18

gatggggctc aggttcat

<210>19

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>19

gatggggctc acgttcat

<210>20

<211>18

<212>DNA

<213>Artificial sequence

444555666777888999

<220>

<223>Sample oligonucleotide

<400>20

gatggggctc atgttcat

<210>21

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>21

gatggggctc gagttcat

<210>22

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>22

gatggggctc gggttcat

<210>23

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>23

gatggggctc gcgttcat

<210>24

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>24

gatggggctc gtgttcat

<210>25

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>25

gatggggctc cagttcat

<210>26

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>26

gatggggctc cggttcat

<210>27

<211>18

1043033600

<223>Sample oligonucleotide

<400>30

gatggggctc tggttcat

<210>31

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>31

gatggggctc tcgttcat

<210>32

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>32

gatggggctc ttgttcat

<210>33

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>33

gatgggcctc aagttcat

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500

<210>34

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>34

gatgggcctc aggttcat

<210>35

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>35

gatgggcctc acgttcat

<210>36

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>36

gatgggcctc atgttcat

<210>37

<211>18

<212>DNA

401611-33460

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>37

gatgggcctc gagttcat

<210>38

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>38

gatgggcctc gggttcat

<210>39

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>39

gatgggcctc gcgttcat

<210>40

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

1.0.3.0 = 23.04.2016

<400>40

gatgggcctc gtgttcat

<210>41

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>41

gatgggcctc cagttcat

<210>42

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>42

gatgggcctc cgtttcat

<210>43

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>43

gatgggcctc ccgtttcat

<210>44

U.S. PATENT & TRADEMARK OFFICE

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>44

gatgggcctc ctgttcat

<210>45

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>45

gatgggcctc tagttcat

<210>46

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>46

gatgggcctc tggttcat

<210>47

<211>18

<212>DNA

<213>Artificial sequence

SEQUENCE SHEET

<220>

<223>Sample oligonucleotide

<400>47

gatgggcctc tcgttcat

<210>48

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>48

gatgggcctc ttgttcat

<210>49

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>49

gatgggtctc aagttcat

<210>50

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>50

Genebank

<223>Sample oligonucleotide

<400>57

gatgggtctc cagttcat

<210>58

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>58

gatgggtctc cggttcat

<210>59

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>59

gatgggtctc ccgttcat

<210>60

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>60

gatgggtctc ctgttcat

Downloaded from www.sagepub.com

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>64

gatgggtctc ttgttcat

<210>65

<211>18

<212>DNA

<213>p53 fragment

<220>

<223>Sample oligonucleotide

<400>65

atgaaccgga ggcccatc

<210>66

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

<400>66

atgaaccaga ggcccatc

<210>67

<211>18

<212>DNA

<213>Artificial sequence

<220>

<223>Sample oligonucleotide

401650334460

<400>67

atgaaccgga gtcccatc

atgaaccgga gtcccatc